GENERATING OPTIMIZED MARKER-BASED RIGID BODIES FOR OPTICAL TRACKING SYSTEMS

Frank Steinicke, Christian Jansen and Klaus Hinrichs Institut für Informatik, WWU Münster, Einsteinstr. 62, 48149 Münster, Germany

Jan Vahrenhold

Informatik XI, Universität Dortmund, 44221 Dortmund, Germany

Bernd Schwald

TWT GmbH, Information & Engineering Technologies, 73765 Neuhausen, Germany

Keywords: Optical tracking, model-based object tracking, rigid bodies.

Abstract: Marker-based optical tracking systems are often used to track objects that are equipped with a certain number of passive or active point markers. Fixed configurations of these markers, so-called rigid bodies, can be detected by, for example, infrared stereo-based camera systems, and their position and orientation can be reconstructed by corresponding tracking algorithms. The main issue in designing the geometrical constellation of these markers and their 3D positions is to allow robust identification and tracking of multiple objects, and this design process is considered to be an essential and challenging task. At present, the design process is based on trial-and-error: the designer constructs a marker configuration, evaluates it in a given setup, and rearranges the marker positions within the configuration if necessary. Even though single ready-made rigid bodies permit sufficiently good tracking, it is not ensured that the corresponding arrangements of markers meet any quality criteria in terms of reliability and robustness. Furthermore, it is unclear whether it is possible to add further rigid bodies to the setup which are sufficiently distinguishable from the given ones.

In this paper, we present an approach to semi-automatically generate marker-based rigid bodies which are optimal with respect to the properties of the tracking system for which they are used, e.g., granularity, accuracy, or jitter. Our procedure which is aimed at supporting the design process as well as improving tracking generates configurations for several devices associated with an arbitrary set of point-based markers. We discuss both the technical background of our approach and the results of an evaluation comparing the tracking quality of commercially available devices to the rigid bodies generated by our approach.

1 INTRODUCTION

The usage of *optical tracking systems* based on infrared (IR) light is becoming more and more common for virtual, augmented or so-called mixed reality (MR) systems used in several application domains. This is due to the fact that these systems provide a large interaction space and quite high accuracy, and in contrast to mechanical approaches no wires disturb the interaction. Furthermore no interferences may occur as when using magnetic or ultrasonic technology. Nowadays IR-based optical tracking systems exist as prototypes in research institutes (Kato and Billinghurst, 1999; Dorfmüller-Ulhaas, 2002; Schwald, 2005; Ribo et al., 2001) and are also com-Steinicke F, Jansen C,, Hinrichs K, Vahrenhold J. and Schwald B. (2007). mercially available (A.R.T., 2006; Fakespace Systems, 2006). The main issue in designing tracking system is to advance both hardware as well as algorithms in order to increase the accuracy and robustness. These factors are the most important properties to make such a system usable for applications, e.g., in medicine or MR environments in general.

When using such systems objects or devices to be tracked are associated with so-called *rigid bodies*, sometimes denoted as *targets* (Schwald, 2005; Kato and Billinghurst, 1999; Dorfmüller-Ulhaas, 2002). A rigid body is a fixed geometrical arrangement of at least three passive or active IR markers. A calibrated camera system allows to reconstruct the 3D coordinates of IR point-based markers in the tracking coor-

GENERATING OPTIMIZED MARKER-BASED RIGID BODIES FOR OPTICAL TRACKING SYSTEMS.

In Proceedings of the Second International Conference on Computer Vision Theory and Applications - IU/MTSV, pages 387-395 Copyright © SciTePress dinate system if at least two cameras can detect the same marker (Zhang, 2000). A rigid body registered by the tracking system, can be tracked by means of the fixed known distances between its markers.

In many MR applications, devices like head mounted displays, stereo glasses or interaction devices, e.g., gloves or wands, have to be tracked (Bowman et al., 2004). Often these devices do not require more than three, four or five markers, arranged in a target attached to such a device. For instance, a typical setup in a co-located MR environment consists of two users, each equipped with an interaction device and stereo glasses each including approximately three to five markers. However, some applications need to track more complicated objects requiring a much larger number of markers. This occurs, for example, when tracking real objects having several markers attached on different sides in order to avoid occlusions.

Usually, rigid bodies are defined by the developer or an interaction device designer by arranging markers around the object respectively device to be tracked (Davis et al., 2004). When arranging markers on such an object it is essential to arrange them in such a way that the distances between all markers are pairwise different if possible. Otherwise the tracking system may mistake distances and the corresponding device will not be recognized. For each new marker to be integrated into a configuration consisting of n markers n new distances have to be considered. For example, when building a simple target including three markers, adding a fourth marker requires the designer to consider three new distances. Moving one marker in a configuration of four markers may change three distances that have to be pairwise different and different from each of the distances of the remaining three markers. Thus, although finding a well-defined configuration seems to be simple it involves a non-trivial task of arranging the markers especially if several devices with numerous markers are included.

However, when constructing a target usually the markers are arranged by trial-and-error. After a prototype rigid body has been built, the application developer or user has to test the corresponding device in a laboratory setup. If the test shows bad *rigid body performance*, the designer has to rebuild the device. A bad rigid body performance means that the device is often not tracked or it is tracked with position or orientation errors which do not result from accuracy errors of the used tracking system; these mistakes result from confusing distances within the same configuration or between different configurations.

After several iterations of building, testing and redefining, the designer may have constructed a configuration that provides sufficient tracking properties. However, even when ready-made targets are tracked well, it is not ensured that the corresponding arrangements of markers are optimal in terms of reliability and robustness and if it is possible to add further targets to the setup that are distinguishable from the already designed ones. Typically, targets built via such a procedure consist of distances which have the potential to disturb the tracking process. For instance, distances within the same target or distances of different targets used for the interaction may be equal.

In order to support the arrangement of markers in a target we present a procedure to semi-automatically generate marker-based rigid bodies in an iterative way. When using our approach the proposed configurations are adapted to the properties of the corresponding tracking system, e.g., granularity, accuracy, jitter etc., and thus the described concepts enhance the tracking process. Our procedure allows to generate rigid bodies for several devices associated with an arbitrary set of markers. This paper describes the technical background of our approach and the results of an evaluation comparing commercially available devices with their associated targets to rigid bodies proposed by our approach.

The remainder of this paper is structured as follows. Section 2 outlines the concepts of optical tracking and explains how 3D points are generated from 2D images of point-based markers grabbed with at least two cameras. In Section 3 we describe how rigid bodies are defined and how the detection of a configuration is performed by the tracking system. Section 4 explains our algorithm to generate configurations semi-automatically in an iterative way. In Section 5 we present an evaluation of our concept and show how we could increase the performance of the devices by redefining two example targets. Section 6 concludes the paper and gives an overview about future research directions.

2 INFRARED-BASED OPTICAL TRACKING

Since the brightness of most MR systems, e.g., CAVEs, PowerWalls, etc., is relatively limited, many projection-based environments require a significant reduction of the ambient light. To overcome the resulting lighting problem for the cameras, an *infrared* (*IR*) optical tracking systems illuminates the scene using infrared light, and IR pass filters are attached to the lenses of the tracking system's cameras. Infrared optical tracking systems aim at measuring the (real-world) positions of numerous markers in the environment. Since active markers such as *light-emitting*

During the tracking process, the relation between the current position and orientation of a given target and the position of its *reference position/orientation* specifies the *rigid body transformation* of the target. This rigid body transformation is applied to a virtual object which is associated with a corresponding target, i.e., the transformation matrix describing the rigid body transformation is applied to the associated scene object visualized in the virtual environment.

To further increase the effectiveness of the tracking algorithm, learning algorithms may be used that aim at enhancing the tracking process, e.g., by tuning the distances predefined manually in the reference target with respect to measured distances (Figueriredo, 2002; Kanbara et al., 2001).

3.1 Description

A target used to track an input device is a configuration of a set of markers where one of the markers serves as a local reference point. As an example, consider the six-degrees-of-freedom interaction device depicted in Figure 3. The target attached to this device consists of three markers where the point $m_1 :=$ (0,0,0) is associated with the marker at the top of the input device, point $m_2 := (-73.0, 0.0, -188.72)$ is associated with the marker in front of the handle, and $m_3 := (0.0, 0.0, -181.25)$ is associated with the marker at the top of the stick branched to the left. The positions of the markers are given as relative coordinates (in millimeters) in the Cartesian coordinate system with the origin at m_1 .



Figure 3: Example device with an associated with a rigid body consisting of three markers.

In general, such targets can be defined by two different procedures: (1) storing the (absolute) positions of the markers or (2) storing the relative distances between the markers as well as its reference position/orientation. We use the second alternative and apply a matrix representation for the distances

$$\mathcal{D}_{\mathcal{M}} = \{ d(m_i, m_j) \mid m_i, m_j \in \mathcal{M}; i, j = 1, ..., n \} \quad (1)$$

where \mathcal{M} is the set of all markers in a configuration consisting of *n* markers, and $d(m_i, m_j)$ denotes the Euclidean distance between two markers m_i and m_j in 3D-space. The quadratic $n \times n$ -dimensional matrix $D = (d_{ij} := d(m_i, m_j))_{i,j=1,...,n}$ stores the distances induced by the configuration and has the following form:

$$D_{\mathcal{M}} = \begin{pmatrix} 0 & d_{12} & d_{13} & \dots & d_{1n} \\ d_{21} & 0 & d_{23} & \dots & d_{2n} \\ d_{31} & d_{32} & 0 & \dots & d_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ d_{n1} & d_{n2} & d_{n3} & \dots & 0 \end{pmatrix}$$
(2)

From (1) and (2), it follows:

1.
$$d(m_i, m_i) = 0$$
, and

2. *d* is symmetric, i.e., $d(m_i, m_j) = d(m_j, m_i)$.

To ensure that no errors occur in determining the correct rigid body transformation, we also desire that $d(m_i, m_j)$ are pairwise different for all $1 \le i < j \le n$.

3.2 Detection

The above matrix-based description of rigid bodies can be used to identify the rigid body by means of *distance detection*. When tracking a target taken from a set of several predefined rigid bodies, the tracking algorithm scans the 3D point cloud \mathcal{P} resulting from the reconstruction of the detected markers. While doing so, the algorithm tries to match point-to-point distances in \mathcal{P} to distances $d_{ij} = d(m_i, m_j)$ stored in the *i*-th row and *j*-th column of each matrix *D* that describes one of the predefined rigid bodies (see (2)).

If any d_{ij} varies by at most a *distance threshold* of ε from some $d \in \mathcal{D}_{\mathcal{M}}$, i.e., if

$$|d - d_{ij}| < \varepsilon, \tag{3}$$

it is assumed that $d = d_{ij}$, and — assuming that all distances are unique — the first two markers of a target T, m_i and m_j , are detected within the point cloud \mathcal{P} . Otherwise the algorithm scans \mathcal{P} for the next distance stored in $\mathcal{D}_{\mathcal{M}}$.

The algorithm continues searching until a third marker m_k has been found or the search space is exhausted. If a third marker m_k is identified by detecting the distances d_{ik} or d_{jk} in the point cloud, the rigid body transformation of this target with respect to its reference position can be determined (Dorfmüller-Ulhaas, 2002). From Equation (3), we see that if we choose ε to small, distances may not be found due to accuracy errors. If the threshold is chosen too high, there may be ambiguities with other distances. This may result in tracking of targets with wrong position/orientation or to confusions between targets. Hence, the threshold must be carefully adapted in order to increase the reliability of the tracked data.

4 GENERATION OF TARGETS

Assuming that the shape and size of the markers are identical for each rigid body we focus on improving the tracking performance by means of redefining positions of markers within a target in order to be able to support the largest possible distance threshold. Hence, the main idea is to arrange the markers in such a way that the resulting distances between each pair of markers are as diverse as possible with respect to the granularity of the tracking system and already existing configurations.

4.1 Iterative Approach

Let $\mathcal{D}_{\mathcal{M}_r}$ denote the set of all distances of the *r*-th rigid body, which consists of the marker set \mathcal{M}_r . Furthermore, let \mathcal{D} denote the union of all $\mathcal{D}_{\mathcal{M}_r}$ for all rigid bodies registered at the tracking system. With increasing difference between the distances contained in \mathcal{D} , the threshold ε from Equation (3) can be chosen larger. We calculate the threshold ε by

$$\varepsilon = \min_{r} \{ \frac{|d_{ij} - d_{kl}|}{2}, d_{ij}, d_{kl} \in \mathcal{D}_{\mathcal{M}_{r}}; d_{ij} \neq d_{kl} \}.$$
(4)

Hence, ε is given by the minimal difference between two distances within a configuration of a rigid body involved in the tracking process. As it can easily be seen from Equation (3) a large threshold improves the tracking robustness, it is beneficial to arrange the markers in such a way that this minimal difference is as large as possible. However, the distances within rigid bodies are constrained by the maximum size of the target with respect to the device to which it is attached; a large target may be inconvenient, heavy and it may restrict the user's degrees of freedom. In general for hand-held devices the maximum size of a target is about 20 - 30cm.

Our approach provides the largest possible threshold with respect to the granularity of the used tracking system. The granularity defines the minimal distance between two points that can be measured. If a point Pis tracked with an accuracy determined by the granularity g, it can be ensured that P is located in a sphere around P with radius g. For this reason, two distances between markers are well-defined only if they distinguish at least by $2 \cdot g$. While in optical tracking systems the granularity is in the area of submillimeters, the sizes of the markers itself measures at least about 4mm, and therefore we approximate the granularity to 8mm upwards.

Assume that the maximal distance in a configuration is predefined by y, the interval of distances [0, y] is decomposed into subintervals with length of $2 \cdot g$. Hence, the set of all distinguishable distances between markers *C* is given by

$$C = \left\{ d_i \mid d_i := 2 \cdot i \cdot g; \ i = 1, ..., \left\lfloor \frac{y}{2 \cdot g} \right\rfloor \right\}.$$
(5)

The number of markers that can be integrated into a tracking system providing that all resulting distances must vary is constrained to $\max\{i \mid \sum_{j=1}^{i} (j-1) < \frac{y}{2 \cdot g} \mid \}$.

4.1.1 Designing a New Target

We start with a tracking system without any registered rigid bodies, i.e., \mathcal{D} is an empty set. In order to define a new target, the developer has to specify the properties of the tracking system, i.e., granularity *g* and maximum distance *y* of the target. Now, our approach supports the designer in deciding which distances from C should be taken into account for a new marker configuration.

Only one distance, namely d_{12} , results from the first two markers m_1 and m_2 . Using our approach the designer can either specify the largest distance $d_{12} \in C$ that should be used in the configuration, or our algorithm starts with $2 \cdot g$ and iterates through C until an optimal configuration is found as described following.

For simplicity, we place m_1 to the origin and m_2 onto the *z*-axis at a distance of d_{12} :

$$m_1 := (0.0, 0.0, 0.0)$$

 $m_2 := (0.0, 0.0, -d_{12})$

Since, \mathcal{D} is the set of all distances which are already included in rigid bodies, using our approach implies $\mathcal{D} \subset C$. After the first distance of the target is determined, we add d_{12} to \mathcal{D} .

As mentioned above, to allow six-degrees-offreedom tracking, at least three markers are required in a fixed configuration. Hence, our approach determines the best position for the marker m_3 , such that the resulting distances $d_{13}, d_{23} \in C$ are as diverse as possible from d_{12} and each other. This is done by choosing d_{13} and d_{23} such that they are uniformly distributed within the subintervalls of C.



Figure 4: Example configuration of used distances from C for a rigid body consisting of three markers.

Figure 4 illustrates this procedure. Let d_1 , d_2 etc. denote the distances from C (see Equation (5)). After the user has specified, for example, $d_{12} := d_7$ as first

distance between m_1 and m_2 , the next marker is inserted in such a way that both new distances $d_{13} := d_1$ and $d_{23} := d_4$ result in the best distribution possible. When determining these markers by means of calculating the corresponding distances, it has to be ensured that the spheres around m_1 with radius d_{13} and around m_2 with radius d_{23} intersect at least in one point which defines the position of marker m_3 . Otherwise new radii, i.e., distances from *C*, have to be tested for intersections. This procedure is illustrated in Figure 5. When more than one intersection point exists, the marker m_3 is chosen to be located in the local *xy*-plane of the configuration having a positive *y*-value.



Figure 5: 2D Illustration of determination of third point using spheres.

Again, we add d_{13} and d_{23} to \mathcal{D} . Now, that three markers are configured, the rigid body can be tracked. In order to enable robust tracking even if several markers are occluded, e.g., by the device itself, further markers can be added.

4.1.2 Adding a Further Marker

We continue with a set of targets each consisting of several markers, where the resulting distances are stored in \mathcal{D} . Let the current target contain i-1 markers. If the user wants to add the *i*-th marker to the target, i - 1 distances from C have to be selected uniformly analogously to the procedure described in Section 4.1.1. Furthermore, the new distances shall be elements of $\mathcal{C} \setminus \mathcal{D}$, i.e., they are used neither within the current target nor within another configuration. Again the spheres surrounding the markers of the current target with radii determined by the corresponding distances have to intersect in one point at least. If one or more intersection points exist, the desired distances can be satisfied and the new marker can be inserted. Otherwise, all distances have to be redefined, i.e., the next distance from C is used as first distance in the configuration and it is continued as described above.

An *error degree* is calculated by means of the sum of the reciprocal values of the squares of differences between the distances in \mathcal{D} . This error degree indicates how well the distances are distributed. The

aforementioned procedure is done in an optimization step until the error degree is minimized, and the best configuration for the target results.

If no intersection point exists for any distance of $C \setminus D$, a compromise has to be accepted and also distances from D are allowed as long as no symmetrical triangle constellations result. Thus, another marker can be added although the resulting distances have the potential to introduce ambiguities when tracking. However, since most marker positions have been added using the described iterative way, our approach provides a sufficient set of well-defined markers that enhances tracking.

4.2 Example Configurations

4.2.1 Haptic Input Device

We have evaluated our approach for a hand-held interaction device, called *haptic input device* that is used for multimodal VR-based interactions in several application domains (Steinicke et al., 2005). The associated target is illustrated in Figure 3 and Figure 6 (a).

The distances $d_{12} = 181.25mm$, $d_{23} = 73.38$, and $d_{13} = 202.35$ result from the marker positions as described in Section 3.1. Since the minimal difference between these distances is about 20mm, we have chosen a distance threshold of $\varepsilon = 9mm$ as explained in Equation (4).



Figure 6: Configuration of markers for (a) original target and (b) target generated by our approach. The markers have been scaled for illustration purposes.

We have used our approach to redefine the position of marker m_3 , because we wanted to provide the same starting distance d_{12} in the resulting configuration. The algorithm produces the target depicted in Figure 6 (b). The new position of marker m_3 is (-17.2, 0.0, -124.1). Now, we can increase the distance threshold ε to 22mm, and the error degree could be decreased to 0.001904 from 0.009276 for the original distances.

4.2.2 Medarpa Display

Since the device in Section 4.2.1 is constrained to only three markers, we have tested the approach also

ments as combination of translations and rotations. In both series we have used a distance threshold of 9mm.

Table 2: Results for original (HID-old) and proposed (HIDnew) target of HID for rotational movements in (A) inner, (B) middle, and (C) outer third of interaction region.

	device:	HID-old	HID-new
Α	abs/rel	326 / 82%	326 / 82%
В	abs/rel	282/71%	307 / 77%
С	abs/rel	231 / 58%	263 / 66%

Table 3: Results for original (HID-old) and proposed (HIDnew) target of HID for complex movements in (A) inner, (B) middle, and (C) outer third of interaction region.

	device:	HID-old	HID-new
Α	abs/rel	318 / 80%	326 / 82%
В	abs/rel	249 / 62%	287 / 72%
С	abs/rel	239 / 60%	274 / 69%

We have performed a similar analysis for the original and proposed target of the MEDARPA display. Therefore, we have taken 5 series with 2000 measurements using a 2.5mm as well as 6mm distance threshold.

While the original target has been correctly tracked 1525 in average (76.5%), our proposed configuration has reached in average 1662 correct tracking events (83.1%) when using the 2.5mm distance threshold. For 6mm distance threshold, the results are even better: 1529 in contrast to 1752 correct tracking events for the original target respectively the proposed configuration, which corresponds to an improvement of 14%.

The results clearly show that our proposed configuration de facto improves the tracking robustness. In all regions of the interaction space, and with all distance thresholds, our proposed configuration performed better than the original rigid bodies.

6 DISCUSSION

In this paper we have proposed an approach to automatically generate rigid bodies for arbitrary MR applications running with optical-based tracking systems. The approach determines the optimal configuration for a target consisting of an arbitrary number of markers with respect to the properties of the used tracking system. We have tested the approach by redefining marker positions for existing devices. An evaluation of the proposed configurations shows the benefits of the approach; improvements of up to 20% could be achieved without any modifications to the tracking system. The considered devices have proven their benefits for many applications in research as well as industrial usage, and they have been revealed having well-defined tracking properties. Nevertheless, our simple approach enhances the tracking for them.

For the future we want to expand our system by exporting a construction plan for the designer in order to improve also the build process. This is due to the fact that during our evaluation we figured out that it is essential that the arrangement of markers allows the cameras to see as many markers as possible simultaneously. A bad construction results in markers occluding themselves, which yields reconstruction errors. Furthermore, the size have the markers have to be considered therefore. Moreover, the used camera setup has to be taken into account since it has a major impact on the tracking performance.

REFERENCES

- A.R.T. (2006). http://www.ar-tracking.de. Advanced Real-Time Tracking.
- Azarbayejani, A. and Pentland, A. (1995). Camera Selfcalibration from one Point Correspondence. Technical Report Perceptual Computing Technical Report 341, MIT Media Laboratory.
- Bowman, D., Kruijff, E., LaViola, J., and Poupyrev, I. (2004). *3D User Interfaces: Theory and Practice*. Addison-Wesley.
- Davis, L. D., Hamza-Lup, F. G., and Rolland, J. P. (2004). A Method for Designing Marker-based Tracking Probes. In 3rd International Symposium on Mixed and Augmented Reality, pages 120–129. IEEE and ACM.
- Dorfmüller-Ulhaas, K. (2002). Optical Tracking From User Motion to 3D Interaction. PhD thesis, Technische Universität Wien.
- Fakespace Systems (2006). http://www.fakespace.com.
- Figueriredo, P. (2002). Automatic Learning and Detection of Point-based Models. Master's thesis, ZGDV, Darmstadt and Minho University, Portugal.
- Hartley, R. and Sturm, P. (1997). Triangulation. *Computer Vision and Image Understanding*, 68(2):146–157.
- Kanbara, M., Yokoya, N., and I, H. T. . (2001). A Stereo Vision-based Augmented Reality Registration with Extendible Tracking of Markers and Natural Features. In Proceedings of Conference on Computer Vision and Pattern Recognition (CVPR2001), pages 1045–1048.
- Kato, H. and Billinghurst, M. (1999). Marker Tracking and HMD Calibration for a Video-based Augmented Reality Conferencing System. In Proceedings of the 2nd International Workshop on Augmented Reality (IWAR), page 85.
- Ribo, M., Pinz, A., and Fuhrmann, A. (2001). A new Optical Tracking System for Virtual and Augmented

Reality Applications. In *Proceedings of the IEEE Instrumentation and Measurement Technical Confer ence*, volume 3, pages 1932–1936.

- Schwald, B. (2005). A Tracking Algorithm for Rigid Point-Based Marker Models. In Proceedings of International Conferences in Central Europe on Computer Graphics, Visualization and Computer Vision, pages 61–62.
- Schwald, B. and Figueiredo, P. (2004). Learning of Rigid Point-Based Marker Models for Tracking with Stereo Camera Systems. In *1. Workshop der GI VR/AR* (*Chemnitz*), pages 23–34.
- Steinicke, F., Ropinski, T., and Hinrichs, K. (2005). Multimodal Interaction Metaphors for Manipulation of Distant Objects in Immersive Virtual Environments. In 13th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision, pages 45–48.
- ZGDV (2005). MEDARPA: MEDical Augmented Reality for PAtients. http://www.medarpa.de.
- Zhang, Z. (2000). A Flexible New Technique for Camera Calibration. *IEEE Transactions on Pattern Analysis* and Machine Intelligence, 22(11):1330–1334.